Sox4 is a master regulator of epithelial-mesenchymal transition by controlling ezh2 expression and epigenetic reprogramming

Tiwari, Neha ; Tiwari, Vijay ; Waldmeier, Lorenz ; Balwierz, Piotr J ; Arnold, Phil ; Pachkov, Mikhail ; Meyer-Schaller, Nathalie ; Schübeler, Dirk ; van Nimwegen, Erik ; Christofori, Gerhard

Abstract: Gene expression profiling has uncovered the transcription factor Sox4 with upregulated activity during TGF- -induced epithelial-mesenchymal transition (EMT) in normal and cancerous breast epithelial cells. Sox4 is indispensable for EMT and cell survival in vitro and for primary tumor growth and metastasis in vivo. Among several EMT-relevant genes, Sox4 directly regulates the expression of Ezh2, encoding the Polycomb group histone methyltransferase that trimethylates histone 3 lysine 27 (H3K27me3) for gene repression. Ablation of Ezh2 expression prevents EMT, whereas forced expression of Ezh2 restores EMT in Sox4-deficient cells. Ezh2-mediated H3K27me3 marks associate with key EMT genes, representing an epigenetic EMT signature that predicts patient survival. Our results identify Sox4 as a master regulator of EMT by governing the expression of the epigenetic modifier Ezh2.

DOI: https://doi.org/10.1016/j.ccr.2013.04.020

Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming

Neha Tiwari[1,4], Vijay K. Tiwari[2,5], Lorenz Waldmeier[1], Piotr J. Balwierz[3], Phil Arnold[3], Mikhail Pachkov[3], Nathalie Meyer-Schaller[1], Dirk Schübeler[2], Erik van Nimwegen[3], and Gerhard Christofori[1*]

[1]Department of Biomedicine, University of Basel, Basel, Switzerland
[2]Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland
[3]Biozentrum, University of Basel, and Swiss Institute of Bioinformatics, Basel, Switzerland
[4]Current address: Institute of Physiological Chemistry, University Medical Center, Johannes Gutenberg University Mainz, Mainz
[5]Current address: Institute for Molecular Biology gGmbH (IMB), Mainz, Germany

*Corresponding author: Gerhard Christofori
Department of Biomedicine
University of Basel
Mattenstrasse 28
4058 Basel
Switzerland
Tel. +41 61 267 35 64
Fax. +41 61 267 35 66

e-mail: Gerhard.christofori@unibas.ch

Running title: Sox4 regulates EMT via Ezh2

Subject Category: Molecular biology of disease, cancer

Keywords: Sox4, TGFβ, EMT, Ezh2, tumorigenesis, metastasis

Word Count: 63'993 characters including spaces (all text, references, figure legends)
8 Figures, Supplemental Information including 6 Supplemental Figures, 2 Supplemental Tables, Supplemental Figure Legends, and Supplemental Experimental Procedures.

Summary

Gene expression profiling has uncovered the transcription factor Sox4 with up-regulated activity during TGFβ-induced EMT in normal and cancerous breast epithelial cells. Sox4 is indispensable for EMT and cell survival *in vitro* and for primary tumor growth and metastasis *in vivo*. Among several EMT-relevant genes, Sox4 directly regulates the expression of *Ezh2*, encoding the Polycomb group histone methyltransferase that trimethylates histone 3 lysine 27 (H3K27me3) for gene repression. Ablation of Ezh2 expression prevents EMT, while forced expression of Ezh2 restores EMT in Sox4-deficient cells. Ezh2-mediated H3K27me3 marks associate with key EMT genes, representing an epigenetic EMT signature that predicts patient survival. Our results identify Sox4 as a master regulator of EMT by governing the expression of the epigenetic modifier Ezh2.

Significance

An epithelial-mesenchymal transition (EMT) is a key process during organismal development and the progression of epithelial tumors to metastatic cancers. Here, we identify the transcription factor Sox4 as a master regulator of EMT. Sox4 appears to act highly upstream in the epistatic hierarchy of EMT regulation. It controls a number of EMT-relevant genes as well as *Ezh2*. Conversely, Ezh2 is able to functionally replace Sox4 during EMT and regulates the expression of a number of EMT-associated genes. *Ezh2* thus represents one critical Sox4 target gene during EMT. The results exemplify an important interplay between transcriptional and epigenetic control during EMT and suggest that the inhibition of Ezh2 could be an attractive avenue for the therapeutic intervention of malignant tumor progression.

Highlights

- Sox4 is critical for EMT and for experimental primary tumor growth and metastasis
- Sox4 directly regulates EMT-relevant genes, among them *Ezh2*
- Ezh2 function and thus H3K27me3 are required for EMT
- The expression of Ezh2-regulated genes is predictive for patient survival

Introduction

Epithelial–mesenchymal transition (EMT) is a cellular mechanism known to constitute the core of normal embryonic development (Hanahan and Weinberg, 2011; Kalluri, 2009; Kalluri and Weinberg, 2009; Nieto, 2010; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006). Similar, yet pathophysiological transitions occur during the progression of epithelial tumors, endowing cancer cells with increased motility and invasiveness to seed metastasis and, sometimes, after metastatic dissemination to re-differentiate into epithelial structures by mesenchymal-epithelial transition (Kang and Massague, 2004; Thiery and Morgan, 2004). Multiple oncogenic events and signaling pathways, mediated for example by transforming growth factor β (TGFβ), hepatocyte growth factor (HGF), Wnt and Notch signaling or by oncogenic Src or Ras activation, are implicated in the induction of EMT (Hanahan and Weinberg, 2011; Kalluri, 2009; Kalluri and Weinberg, 2009; Nieto, 2010; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006).

Sox4 is a member of the Sox (SRY-related HMG-box) family of transcription factors with a critical role in embryonic development and in cell-fate determination during organogenesis of the heart (Restivo et al., 2006; Schilham et al., 1996), pancreas (Lioubinski et al., 2003; Wilson et al., 2005) and brain (Cheung et al., 2000; Hong and Saint-Jeannet, 2005), and in B and T lymphocyte differentiation (Cheung et al., 2000; Hong and Saint-Jeannet, 2005; Lioubinski et al., 2003; Schilham et al., 1997; Schilham et al., 1996; van de Wetering et al., 1993; Wilson et al., 2005). Sox4 gene expression is up-regulated in many cancer types, and increased Sox4 activity contributes to cellular transformation (Liu et al., 2006; Shin et al., 2004), cell survival (Aaboe et al., 2006; Ahn et al., 2002; Liu et al., 2006; Pramoonjago et al., 2006) and metastasis (Liao et al., 2008; Tavazoie et al., 2008). For example, restoration of miR-335 expression suppresses lung and bone metastasis in human cancer cells by interfering with the expression of Sox4 (Tavazoie et al., 2008). Interestingly, Sox4 also directly modulates key cellular regulators, including the genes encoding for epidermal growth factor receptor (EGFR), tenascin-C, heat shock protein 70 (Hsp70), frizzled homolog 5 (Fzd5), Delta-like 1 (Dll1), Patched-1 (Ptch1), the transcriptional regulators Mll, Foxa1, Znf281 and Nkx3-1, and components of the RNAi machinery, including Dicer, Argonaute 1 and RNA helicase A (Scharer et al., 2009). Finally, Sox4 regulates Wnt signaling by directly binding to β-catenin and Tcf family members (Sinner et al., 2007). Most recently, Sox4 has been reported to induce EMT and to cooperate with oncogenic Ras in breast cancer progression (Zhang et al., 2012), however, its direct transcriptional target genes during EMT have remained elusive. Here, we demonstrate a central

role of Sox4 in EMT as well as in primary tumor growth and metastasis by directly regulating the expression of a number of genes with critical functions in EMT, including *Ezh2*.

The histone methyltransferase Ezh2 (Enhancer of Zeste homolog 2) is a component of the Polycomb (PcG) repressive complex 2 (PRC2) which epigenetically regulates genes involved in cell fate decisions. Ezh2 specifically trimethylates nucleosomal histone H3 at lysine 27 (H3K27me3), an epigenetic modification associated with gene silencing (Sparmann and van Lohuizen, 2006). A conditional knockout of the *Ezh2* gene in basal keratinocytes leads a precocious acquisition of an epidermal barrier function in the embryo (Ezhkova et al., 2009), while a conditional knockout of *Ezh2* in B-lymphocytes results in improper IgH re-arrangements (Su et al., 2003), suggesting a role for Ezh2 methyltransferase in cell differentiation and maturation. Ezh2 is found highly expressed in a variety of cancer types where the genomic loss of miR-101 leads to increased expression of Ezh2 and concomitant deregulation of epigenetic pathways, altogether resulting in cancer progression (Varambally et al., 2008). Furthermore, Ezh2 can induce EMT and increase the metastatic potential of prostate cancer cells by down-regulation of DAB2IP, a tumor-suppressive Ras GTPase activating protein (RasGAP) (Chen et al., 2005; Min et al., 2010). Finally, low Ezh2 expression levels correlate with metastasis-free survival in breast cancer (Cao et al., 2008; Kleer et al., 2003). However, whether and how Ezh2-mediated epigenetic mechanisms contribute to the transcriptional reprogramming that accompanies EMT is still poorly understood.

Results

Sox4 is required for EMT

To identify critical genes underlying early, intermediate and late stages of EMT, we induced EMT in the untransformed normal murine mammary gland cell line NMuMG by treatment with TGFβ for 0, 1, 4, 7, 10 and 20 days (data not shown). During this time course, the cells underwent progressive EMT and acquired a complete mesenchymal morphology (Lehembre et al., 2008). Motif Activity Response Analysis [MARA; Suzuki et al., 2009)] of gene expression data derived from the EMT time course predicted several transcription factor binding motifs to be important regulators of the EMT expression dynamics, including a motif bound by Sox transcription factors (Figure S1A). Gene expression profiling and quantitative RT-PCR experiments revealed that Sox4 and 9 were the only Sox family members significantly up-regulated in their expression during TGFβ-induced EMT in NMuMG cells (Figure S1B-D). Sox4 expression was also robustly induced during EMT in a number of other cellular EMT systems, including TGFβ-induced EMT in Py2T murine breast cancer cells derived from a tumor of

MMTV-PyMT transgenic mice (Waldmeier et al., 2012), in MCF10A human breast epithelial cells, in EpRas murine mammary epithelial cells (Figure S1D-G), and in HMEC human mammary epithelial cells (Kloeker et al., 2004). Sox4 expression was also increased in MCF7 human breast cancer cells that exert full EMT upon shRNA-mediated knockdown of E-cadherin expression (Lehembre et al., 2008; Figure S1H).

To directly assess the role of Sox4 and Sox9 in EMT, NMuMG cells were transfected with a pool of two different siRNAs against Sox4 (siSox4) and Sox9 (siSox9), as well as siSox4 and siSox9 together, to transiently ablate the expression of Sox4 and Sox9 in the absence and presence of TGFβ, resulting in an efficient repression of Sox4 and Sox9 expression (Figure 1A; Figure S1I, J). As reported recently (Zhang et al., 2012), Sox4-ablated NMuMG cells were not able to undergo EMT and largely retained their epithelial morphology during TGFβ treatment, while cells transfected with control-siRNA (siControl) changed to a mesenchymal, fibroblastoid phenotype (Figure 1B). In contrast, the single ablation of Sox9 did not have any effect on TGFβ-induced EMT, while the concomitant ablation of Sox4 and Sox9 mimicked the single ablation of Sox4 (Figure S1K, L). Immunofluorescence staining revealed that Sox4 depletion in the absence of TGFβ did not substantially affect the epithelial morphology of NMuMG cells (Figure 1C). In the presence of TGFβ, Sox4-depleted cells maintained the epithelial markers E-cadherin and ZO-1 at the cell membrane, and the expression of the mesenchymal marker N-cadherin was decreased as compared to TGFβ-treated siControl cells (Figure 1C). Moreover, TGFβ-mediated remodeling of the cytoskeleton from cortical actin to stress fibers, as determined by phalloidin staining, and focal adhesion formation, as determined by paxillin staining, were significantly reduced in siSox4 cells (Figure 1C, Figure S1M). The failure to lose epithelial and to gain mesenchymal marker expression in TGFβ-treated, Sox4-depleted cells was further confirmed by immunoblotting analysis (Figure 1D). Notably, Sox4 depletion in NMuMG cells treated previously with TGFβ for 15 days reverted these mesenchymal cells to an epithelial phenotype (Figure 1E). Similar to NMuMG cells, shRNA-mediated stable knockdown of Sox4 in Py2T cells markedly delayed TGFβ-induced EMT (Figure S1N-P). Together, these results demonstrate that Sox4 but not Sox9 is required for the induction and maintenance of TGFβ-induced EMT in murine mammary epithelial cells.

Sox4 is required and sufficient for cell survival and cell migration

Sox4 has been implicated in cell survival in a variety of cancers (Hur et al., 2010; Pramoonjago et al., 2006; Shen et al., 2010). Indeed, Sox4-depleted NMuMG and Py2T cells also showed a

significant reduction in cell growth in comparison to control cells and displayed a discernable sensitivity towards TGFβ-mediated growth inhibition (Figure 2A, B). Propidium iodide staining followed by flow cytometry-based cell cycle analysis revealed a G0/G1 cell cycle arrest after Sox4 depletion (Figure 2C, D). Annexin-V staining demonstrated a marked increase in the levels of apoptosis in siSox4-transfected NMuMG cells (Figure 2E), but not in Py2T cells (data not shown). These results indicate that Sox4 is required for cell survival and proliferation during TGFβ-induced EMT and that the pro-survival function of Sox4 only becomes critical in non-transformed cells. Cell migration is a characteristic feature of EMT and metastasis (Brabletz et al., 2005; Christofori, 2006; Grunert et al., 2003; Huber et al., 2005; Thiery and Sleeman, 2006). Transwell migration assays with the apoptosis-resistant Py2T breast cancer cells revealed that shSox4 Py2T cells migrated significantly less than shControl cells after long-term treatment with TGFβ (Figure 2F).

We next asked whether Sox4 was sufficient to induce EMT. Transient expression of HA-tagged Sox4 in NMuMG cells at the levels observed during TGFβ-induced EMT increased filopodia formation and cell scattering (Figure 3A) accompanied by the gain of the mesenchymal markers N-cadherin and fibronectin (Figure 3B, C). While its mRNA and protein levels were not apparently affected, E-cadherin was displaced at the cell junctions of HA-Sox4 expressing cells (Figure 3D). Stable expression of HA-tagged Sox4 in Py2T cells also induced mesenchymal cell morphology in the absence of TGFβ with a moderate loss of E-cadherin expression but a marked gain in N-cadherin and fibronectin expression (Figure 3E-G). Furthermore, the forced expression of Sox4 promoted Py2T cell migration (Figure 3H). These results suggest that Sox4 is not only required but also sufficient for the induction of an EMT.

Sox4 is required for tumorigenesis and metastatic spread

The critical role of Sox4 in EMT raised the question whether Sox4 contributes to malignant tumor progression and metastasis. First, computational analysis revealed a significant correlation between high Sox4 expression and poor metastasis-free survival in the “Schmidt” database of 200 early-stage, lymph node-negative breast cancer patients (Schmidt et al., 2008) (Figure 4A). The analysis of databases assembled of late-stage breast cancer patient data did not reveal a significant prognostic value for Sox4 expression (data not shown), suggesting that Sox4 may play a critical role in the early stages of the malignant progression of breast cancer.

Next, we investigated the functional contribution of Sox4 to primary tumor growth and metastasis formation *in vivo*. Py2T cells stably expressing an shRNA against Sox4 (shSox4) and control-transfected cells (shControl), both expressing the firefly luciferase gene, were

orthotopically implanted into mammary fat pads of nude mice, and tumor growth and metastasis in lymph-nodes, lungs and livers were quantified. Quantitative RT-PCR (Figure S2A) and immunostaining of histological tumor sections (Figure 4B) documented an efficient depletion of Sox4 expression in the cultured cell line before implantation and in primary tumors, respectively. Ablation of Sox4 expression in Py2T cells lead to a significant reduction in primary tumor growth (Figure 4C). Luciferase activity levels representing the presence of tumor cells in distant organs were found decreased in axillary and inguinal lymph nodes and in lungs and livers of mice transplanted with shSox4 Py2T cells as compared to shControl Py2T cells (Figure S2B), even when the luciferase activities were normalized to the decreased primary tumor sizes observed with shSox4 Py2T cells (metastatic index; Figure 4D). The reduced ability of shSox4 Py2T cells to grow as primary tumors was also observed upon subcutaneous implantation into nude mice (Figure 4E).

We next addressed whether Sox4 is required for metastasis formation in other cancer types. Similar to NMuMG cells, in B16-F10 melanoma cells Sox4 expression was increased upon TGFβ treatment (Figure S2C). Sox4 expression was then ablated in B16-F10 melanoma cells by stable expression of shRNA targeting Sox4 (shSox4; Figure S2D), and the cells were injected into the tail vein of C57Bl/6 mice. Quantification of lung metastasis showed that Sox4-depleted B16-F10 cells were significantly impaired in lung colonization as compared to shControl-expressing B16-F10 cells (Figure S2E, F). Taken together, these results indicate that the transcription factor Sox4 exerts a critical function in primary tumor growth and metastasis formation.

Ezh2 is a direct transcriptional target of Sox4

The critical role of Sox4 in EMT and tumor progression motivated us to identify the genes that were directly regulated by Sox4. We first compared the genome-wide gene expression profiles of siControl and siSox4 NMuMG cells in the absence and presence of TGFβ. Genes found to be differentially expressed in dependence on Sox4 function were further analyzed for Sox4 binding motifs within one kilobase of their transcription start sites (Figure S1A). Of the 189 genes fulfilling these criteria (Table S1), 106 also changed in their expression levels during TGFβ-induced EMT. Gene ontology analysis revealed that 32 of these genes have been previously implicated in processes relevant to tumor progression (Table S2). In order to identify the genes directly regulated by Sox4, we assessed the occupancy of Sox4 at the promoters of these 32 genes. NMuMG cells transiently expressing a HA-tagged version of Sox4 (HA-Sox4) were subjected to Chromatin Immunoprecipitation (ChIP) with HA-specific antibodies followed by

quantitative PCR using primers specific for the region spanning the Sox4 motif found in the promoter of these genes. Out of the 32 genes analyzed, the promoters of 16 and 12 genes were directly bound by Sox4 in NMuMG and Py2T cells respectively, including the key EMT genes *Spred1*, *Edn1*, *Palld*, *Cyr61*, *Ereg*, and *Areg* (Figure S3A - H and data not shown). Genes that were not immunoprecipitated by HA-antibody, as well as an intergenic region, served as negative controls (Figure S3I).

Snail, Zeb and Twist family transcriptional repressors of E-cadherin gene expression are known to play critical roles in the induction of EMT. Hence, we assessed whether Sox4 is epistatic to the expression of these EMT inducers or whether they are regulating Sox4 expression during EMT. Ablation of Sox4 expression in NMuMG cells resulted in the reduced expression of Snail2, Zeb2 and Twist1 but not Snail1 and Zeb1, while in Py2T cells the expression of all five repressors was reduced upon Sox4-depletion (Figure S3J-N and P-T). Conversely, when the five repressors were transiently expressed in NMuMG and Py2T cells, the levels of Sox4 mRNA were not affected (Figure S3O and U). Together, the data indicate that Sox4 acts upstream of the EMT inducers during EMT, most likely in an indirect manner, since none of the EMT inducers have been detected as a direct transcriptional target of Sox4.

Interestingly, the Sox4-dependent gene expression profiling and ChIP experiments also revealed the promoter of the *Ezh2* gene to be directly bound and regulated by Sox4 in NMuMG and Py2T cells (Figure 5A, B). Quantitative RT-PCR and immunoblotting showed that Sox4-depletion lead to significantly reduced Ezh2 expression and a global reduction in H3K27me3 levels in the presence of TGFβ in NMuMG and Py2T cells (Figure 5C). Moreover, the forced expression of Sox4 moderately increased Ezh2 mRNA and protein levels (Figure 3B, C, G). Finally, *Ezh2* promoter activity was reduced upon Sox4 depletion in NMuMG cells in the presence of TGFβ, as determined by *Ezh2* promoter-luciferase reporter assay (Figure 5D). These results indicate that Sox4 is required for *Ezh2* expression during EMT. Yet, it should be noted that *Ezh2* expression levels were already substantial in epithelial cells, where Sox4 expression was low and did not markedly change with the increasing levels of Sox4 during TGFβ-induced EMT (Figure 5E, F; Figure 6B). Hence, Sox4 is required for the efficient expression of Ezh2 during TGFβ-induced EMT, but not in epithelial cells in the absence of TGFβ, suggesting that factors other than Sox4 are critical for *Ezh2* gene expression in the epithelial state of the cells. Consistent with this notion, the dependence of *Ezh2* expression on Sox4 became already apparent between 2 and 8 hours of TGFβ stimulation in NMuMG and Py2T cells (Figure 5E and F).

Loss of Ezh2 function impairs EMT and metastasis

Based on the pleiotropic functions of Ezh2 in mediating the H3K27me3 repressive mark, we speculated whether a major aspect of EMT regulation by Sox4 was dependent on its regulation of *Ezh2* gene expression. Indeed, similar to the ablation of Sox4, siRNA-mediated knockdown of Ezh2 (siEzh2) in NMuMG cells lead to an efficient downregulation of Ezh2 and retention of the epithelial phenotype of NMuMG cells upon TGFβ stimulation (Figure 6A, Figure S4A). Immunoblotting and immunofluorescence microscopy analysis demonstrated a failure to lose E-cadherin and ZO-1 expression as well as a reduced gain in the expression of fibronectin and an impaired formation of actin stress fibers and focal adhesions in Ezh2-depleted cells as compared to siControl-treated cells (Figure 6B and C). These effects became even more apparent in NMuMG cells when both methyltransferases, Ezh1 and Ezh2, were concomitantly depleted (Figure 6A, C and Figure S4B-F). Similar to NMuMG cells, shRNA-mediated stable co-depletion of Ezh1 and 2 in Py2T cells also delayed EMT (Figure S4G-I). Comparable to the ablation of Sox4 expression, Ezh2-depleted NMuMG cells showed significantly increased apoptosis, a block in G0/G1, and a substantial attenuation in their proliferation during TGFβ-induced EMT (Figure 6D-F). Finally, similar to the depletion of Sox4 in B16-F10 melanoma cells, shRNA-mediated ablation of Ezh2 or Ezh1/2 together resulted in a significantly reduced ability of these cells to form lung metastasis upon intravenous injection (Figure S4J). Notably, the concomitant high expression of Ezh2 and Sox4 significantly correlated with poor metastasis-free survival of early-stage, lymph node-negative breast cancer patients (Schmidt et al., 2008) (Figure 6G).

To directly assess whether *Ezh2* is Sox4's major transcriptional target required for its EMT-inducing function, we transiently ablated Sox4 in NMuMG and Py2T cells and concomitantly expressed Ezh2 by transient transfection of an Ezh2 expression plasmid. Notably, the forced expression of Ezh2 overcame the reduced expression of endogenous Ezh2 upon Sox4 depletion and restored TGFβ-induced EMT and changes in marker expression that otherwise were repressed by Sox4 ablation in NMuMG and in Py2T cells (Figure 7A-C; Figure S5A, B). These results indicate that *Ezh2* is one critical target of Sox4 transcriptional regulation during EMT and metastasis formation.

Ezh2 regulates EMT genes via H3K27me3 modification of their promoters

To identify the target genes that are modified by Ezh2-mediated H3K27me3 epigenetic imprint and that are regulated in their expression during EMT, we performed chromatin immunoprecipitation using a H3K27me3-specific antibody upon TGFβ treatment in NMuMG

cells for 0, 1, 4, 7, 10 and 20 days followed by next generation sequencing (ChIP-Seq) in combination with gene expression profiling. Whole genome scanning revealed 970 regions corresponding to 301 genes with high variance in H3K27me3 levels during EMT and high reproducibility in the biological replicates (Figure 8A). Among the 301 genes, genome-wide gene expression analysis identified 46 genes that were transcriptionally up-regulated during EMT and lost their H3K27me3 marks and 3 genes that were transcriptionally down-regulated during EMT and gained H3K27me3 marks (Figure S6A). We validated the changes in H3K27me3 levels at target gene promoters such as *Mcam*, *Pdgfrb*, *Itga5*, *Col4a1*, and *St6galnac4* in NMuMG and Py2T cells (Figure 8B, C and Figure S6B) by ChIP followed by qPCR for the promoter regions of the genes. The expression of these target genes appeared highly dependent on the presence of Ezh1/2 and the H3K27me3 repressive imprint. For example, the promoters of the *Mcam*, *Pdgfrb* and *Itag5* genes were found highly enriched in H3K27me3 marks in untreated NMuMG cells. Depletion of Ezh1/2 led to a loss of global H3K27me3 levels (Figure S6C) and of H3K27me3 at the promoters of the *Mcam*, *Pdgfrb* and *Itag5* genes (Figure 8D, Figure S6D) with a concomitant increase in their expression (Figure 8E-F, Figure S6E, F) in NMuMG and Py2T cells. Conversely, the *Sgsh* gene gained H3K27me3 at its promoter during TGFβ-induced EMT with a concomitant repression of its expression and, upon Ezh1/2 depletion, the lack of H3K27me3 marks on the *Sgsh* promoter during TGFβ-treatment resulted in increased *Sgsh* gene expression (Figure 8G). In addition, we also validated a subset of other H3K27me3-enriched genes that were changing in their expression profile upon depletion of Ezh1/2 (Figure 8H). Gene ontology analysis revealed a prominent enrichment for gene functions implicated in EMT, malignant tumor progression and metastasis (Figure S6G). Furthermore, analysis of the “Minn” database of breast cancer (Minn et al., 2005) revealed that the 49 genes signature correlated with reduced bone metastasis-free survival of patients in highly malignant breast cancers lacking estrogen receptor expression (ER-) or lacking estrogen receptor, progesterone receptor and ErbB2 expression (triple-negative; TN; Figure S6H-K). Moreover, in the “Schmidt” database of early stage breast cancer (Schmidt et al., 2008), the 49 genes signature correlated with the probability to develop distant metastasis (Figure S6L). In line with these findings, transwell migration assays revealed significantly lower chemotactic migration of NMuMG and Py2T cells depleted of both Ezh1 and 2 (Ezh1/2) as compared to siControl-transfected cells upon treatment with TGFβ (Figure S6M). Together, these data indicate that Sox4 regulates *Ezh2* gene expression and that Ezh2-mediated H3K27me3 marks are critically involved in the transcriptomic reprogramming of cells undergoing EMT and that an epigenetic EMT gene signature correlates with breast cancer patient survival.

Discussion

EMT is accompanied by massive changes in cell morphology and behavior, and transcription factors play a pivotal role in controlling the various cellular functions during EMT, such as cell proliferation, cell survival, cell differentiation, cell migration and cell adhesion. (Hanahan and Weinberg, 2011; Kalluri, 2009; Kalluri and Weinberg, 2009; Nieto, 2010; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006). Here, we have identified the transcription factor Sox4 to be up-regulated in its expression and transcriptional activity during EMT in a number of murine and human non-transformed mammary gland epithelial cells and breast cancer cells. The upregulated expression of Sox4 during TGFβ-induced EMT in NMuMG cells was independent of canonical TGFβ signaling, since shRNA-mediated ablation of Smad4 did not substantially affect Sox4 expression (data not shown). Moreover, treatment of NMuMG and Py2T cells with inhibitors against various signaling pathways revealed that the inhibition of TGFβ receptor and Notch signaling and the activation of Wnt signaling interfered with TGFβ-induced Sox4 expression, while the inhibition of EGF receptor increased Sox4 mRNA levels (data not shown). These observations suggest a complex gene regulatory network that drives the increased expression of Sox4 during the early phases of TGFβ-induced EMT, perspectives that warrant further investigations.

We show that Sox4 is not only required for the initiation of EMT, but also for its maintenance. Furthermore, Sox4 is crucial for primary tumor growth and metastatic spread of Py2T breast cancer and B16-F10 melanoma cells. Finally, we find a significant positive correlation between increased Sox4 expression and the metastatic potential of early-stage, lymph node-negative breast cancer in patients. These results are consistent with a recent report demonstrating that Sox4 is critical for an EMT and tumor growth of human breast cancer cells and that Sox4 correlates with poor prognosis in cancer patients (Zhang et al., 2012).

Despite the recent implications of Sox4 in EMT, its direct transcriptional target genes have remained elusive. Here, we show that Sox4 exerts its central function in EMT and tumor progression by directly regulating the expression of a number of genes implicated in EMT, cell cycle regulation, cell survival and cell migration. Notably, one of the direct targets of Sox4 transcriptional control is *Ezh2*, and ablation of Ezh2 function phenocopies the loss of Sox4 during TGFβ-induced EMT. Moreover, the forced expression of Ezh2 overcomes the failure of Sox4-deficient cells to undergo EMT. These results indicate that Ezh2 is a critical downstream effector of Sox4 and that Sox4 mediates epigenetic reprogramming via regulating the expression of epigenetic modulators during EMT. Consistent with this notion, the concomitant

high expression of Sox4 and Ezh2 correlates with poor prognosis in a subset of breast cancer patients. These findings are indeed highly relevant, since they unravel a combinatorial action of key transcription factors and epigenetic regulators in driving the transcriptional reprogramming underlying cell-fate changes during morphogenic processes such as EMT.

The methyltransferase Ezh2 is the best-studied component of the Polycomb Repressor Complex 2 (PRC2). Increased expression or activity of Ezh2 is a marker of advanced and metastatic disease in many solid tumors, including prostate and breast cancer (Chase and Cross, 2011). Our genome-wide study of Ezh2 associated H3K27me3 mark during EMT revealed dynamic changes at a number of genes, of which a subset also shows the corresponding changes in transcriptional state. We further show that Ezh1 and Ezh2 together determine global as well as promoter-specific H3K27me3 levels. This activity modulates the transcriptional state of target genes, including a number of EMT-relevant genes, and it is a prerequisite for EMT. Notably, this epigenetic EMT signature significantly correlates with poor clinical outcome in patients with highly aggressive and metastatic ER- and triple-negative breast cancer. Thus, Polycomb-mediated transcriptional regulation is a critical contributor to the maintenance of epithelial differentiation and to the transcriptomic reprogramming during EMT.

Previously, single gene studies have suggested a role for Ezh2 in EMT regulation. For example, the PRC2 complex is recruited to the promoter of the *Cdh1* gene for its repression (Cao et al., 2008); (Rhodes et al., 2003), and Snail1-mediated repression of *Cdh1* during stem cell differentiation is PRC2-dependent (Herranz et al., 2008). Notably, Ezh2 forms a co-repressor complex with HDAC1, HDAC2 and Snail to repress *Cdh1* expression and to promote nasopharyngeal carcinoma malignancy (Tong et al., 2012). Moreover, the loss of integrin β4 expression during EMT correlates with a decrease in the activating histone modifications H3K9Ac and H3K4me3 and an increase in the repressive histone modification H3K27me3 at its promoter (Yang et al., 2009). Thus, previous single genes studies have clearly demonstrated a role of PcG-mediated mechanisms in carcinogenesis (Mills, 2010; Sauvageau and Sauvageau, 2010; Sparmann and van Lohuizen, 2006). By outlining genome-wide targets of the PcG machinery during the stepwise progression of EMT and by showing a direct role of Ezh1 and 2 in the transcriptional regulation of key EMT genes, our study establishes a widespread and crucial role for PcG in EMT. We show that the transcription factor Sox4 acts upstream of epigenetic reprogramming events during EMT by directly regulating the expression of the chromatin regulator Ezh2. Once expressed, Ezh2 and its accompanying chromatin regulators need to be recruited to specific target genes during EMT. It is surprising that we have not identified the EMT inducers Snail1/2, Twist1/2 and Zeb1/2 as direct targets of PcG-mediated

regulation or of Sox4 transcriptional control. Yet, given the critical role of Snail1/2, Twist1/2 and Zeb1/2 in regulating EMT and in associating with chromatin regulators, it will be important to further delineate the crosstalk of these proteins with the epigenetic machinery and to assess whether and how they are targeting epigenetic regulators to specific target genes.

Sox4 has been previously linked to the etiology of certain cancer types (Penzo-Mendez, 2010). Interestingly, the closely related family member Sox2 is one of the transcriptional factors required to reprogram somatic cells into induced pluripotent stem (iPS) cells (Takahashi and Yamanaka, 2006). In glioma-initiating cells, a cross-talk between Sox4 and Sox2 has been implicated in maintaining the cells' tumorigenicity by Sox4 binding to the *Sox2* gene enhancer region (Ikushima et al., 2009). Together with previous reports that EMT increases the tumor-initiating (cancer stem cell) potential of cancer cells (Mani et al., 2008; Morel et al., 2008), our observation that Sox4 is required for EMT and for metastasis make Sox4 an exciting target for further investigations. Our experimental results indicate that Sox4 exerts its critical regulatory functions upstream of the Snail, Zeb and Twist family transcriptional inducers of EMT, however, without directly affecting their expression. The promoters of the known EMT regulators such as Snail1/2, Zeb1/2 or Twist1/2 also did now show any changes in H3K27me3 levels during EMT and were not affected following Ezh1/2 depletion (data not shown), suggesting that different regulatory pathways may underlie their expression dynamics during EMT. This suggests that independent regulatory pathways converge to control genes that drive EMT. These findings support our efforts to comprehensively understand the gene regulatory networks that underlie an EMT. A recent report demonstrated that Sox9 cooperates with Snail2 to determine mammary stem cells and to promote malignant tumor progression (Guo et al., 2012), yet our study did not find a requirement of Sox9 for TGFβ-induced EMT in NMuMG and Py2T cells. In contrast, another Sox family member, Sox3, was shown to counteract the expression of the EMT-inducer Snail1 and, conversely, Snail1 represses Sox3 gene expression during gastrulation and in human breast cancer cells (Acloque et al., 2011). We can only speculate that the differences in these findings may be based on the functional variety of Sox family members, which are classified into subgroups by the differences in their protein moieties outside of the HMG DNA binding domains and their varying biological functions (Chew and Gallo, 2009).

In summary, the data presented here provide critical insights into the regulatory networks by which one transcription factor, Sox4, regulates an important cell-fate determination event, namely EMT. Notably, the observation that Sox4 regulates a number of EMT-relevant genes and *Ezh2*, and that Ezh2 modifies the expression of a number of genes known to be critical for EMT exemplifies a critical interplay between transcriptional and epigenetic control

during EMT. Our data suggest that the inhibition of Ezh2 function could be an attractive alternative for therapeutic intervention during malignant tumor progression. Indeed, early results with the Ezh2 inhibitor 3-deazaneplanocin (DZNep) are encouraging (Crea et al., 2011), and first results from clinical trials are impatiently awaited for.

Experimental Procedures

For more details see Supplemental Information

Cell lines

A subclone of NMuMG cells (NMuMG/E9; hereafter NMuMG) and MCF7 shControl and MCF7-shEcad have been previously described (Lehembre et al., 2008; Maeda et al., 2005). B16-F10 melanoma cells, EpRas, and MCF10A cells were commercially available. Py2T cells were derived from a breast tumor of MMTV-PyMT transgenic mice (Waldmeier et al., 2012).

Chromatin immunoprecipitation (ChIP)

ChIP experiments were performed as previously described (Weber et al, 2007). In brief, crosslinked chromatin was sonicated to achieve an average fragment size of 500bp. Starting with 100µg of chromatin and 5µg of anti-HA antibody or anti-H3K27me3 antibody. 1µl of ChIP material and 1µl of input material were used for quantitative real-time PCR using specific primers covering the motif of Sox4 in the promoter of target genes or covering the 1000bp promoter region from the transcription start site. Primers covering an intergenic region were used as a control. The efficiencies of PCR amplification were normalized to the PCR product of the intergenic region. Primer sequences are listed in Supplemental Information.

Next Generation Sequencing Analysis

The reads were mapped to the mouse genome assembly NCBI37/mm9 using the bowtie algorithm (Langmead et al., 2009) allowing up to 2 mismatches within the 36bp-long sequence. Reads which mapped to more than 100 loci were discarded from the subsequent analysis. In total, 24 to 34 million reads were mapped to the genome in the first replicate and from 95 to 106 million reads in the second replicate. The mean DNA fragment size was estimated from the mapping profile on both chromosome strands and the reads were shifted by the half of the size towards the middle of the fragment. This data was exported to wiggle tracks for manual inspection in the UCSC genome browser and served as an input for the enrichment analysis. The genome was tiled with 2kb windows overlapping by 1kb and each window was summarized

by the numbers of reads in both replicates at every stage and the number of reads in the input DNA. The H3K27me3-enrichment was calculated assuming binomial statistics of the read counts $z = \frac{\sum_{i=1}^{S} \frac{f_i}{S} - f_b}{\sqrt{\sum_{i=1}^{S} \frac{f_i \cdot (1-f_i)}{S^2 N_i} + \frac{f_b \cdot (1-f_b)}{N_b}}}$, where $f_i = \frac{n_i}{N_i}$ is the frequency in the 2kb window calculated from the number of reads in the window n_i and the total number of reads N_i in the sample *i* and *S*=12 is the number of samples in both replicates. The background frequencies f_b are calculated analogously. Neighboring windows with $z \geq 5$ and $n_b \leq 50$ were merged and intersected with RefSeq loci extended by 1kb upstream and 1kb downstream. From the list of 96'224 such enriched regions a subset was selected by requiring dynamic (mean variance log-frequencies > 0.2) and consistent (Pearson correlation coefficient of log-frequencies > 0.5) methylation changes. The resulting 970 regions lie within 1 kilobase of the loci of 301 genes.

Animal experimentation

All studies involving mice have been approved by the Swiss Federal Veterinary Office (SFVO) and the regulations of the Cantonal Veterinary Office of Basel Stadt (licences 1878, 1907, 1908).

Statistical analysis

Statistical analysis and graphs were generated using the GraphPad Prism software (GraphPad Software Inc, San Diego, CA). All statistical analysis was done by unpaired/paired, two-sided t-test.

Accesssion numbers

Gene expression data of Sox4 knockdown cells in the presence and absence of TGFβ in NMuMG cells and ChIP-seq data of H3K27me3 imprints during TGFβ-induced EMT in NMuMG cells are deposited at Gene Expression Omnibus (GSE44050 and GSE45579, respectively).

Supplemental Information

Supplemental Information including 6 figures, 2 tables, and extended experimental procedures can be found online at XXXXX.

Acknowledgements

We are grateful to H. Antoniadis, I. Galm and P. Schmidt for technical support, and A. Fantozzi for critical comments. We thank I. Nissen (Laboratory of Quantitative Genomics at D-BSSE, ETH Zürich) and S. Dessus-Babus (FMI, Basel) for Illumina sequencing. We thank A.P. Tsou (National Yang-Ming University, Taipeh), A.H. Peters (FMI, Basel), N. Goldfinger and V. Rotter (Weizmann Institute, Rehovot), and G. Berx (VIB, Ghent) for providing important reagents, and P. Demougin for gene expression profiling. This research has been supported by the Swiss National Science Foundation, the Swiss Initiative for Systems Biology (RTD Cellplasticity), the EU-FP7 framework program TuMIC 2008-201662, and the Swiss Cancer League. Research in the laboratory of V.T. is supported by the Deutsche Forschungsgemeinschaft (DFG) grant TI 799/1-1, EpiGeneSys RISE1 program, Marie Curie CIG 322210 and Wilhelm Sander Stiftung 2012.009.1. Research in the laboratory of D.S. is supported by the Novartis Research Foundation, the European Union (NoE ‘‘EpiGeneSys’’ FP7- HEALTH-2010-257082 and the ‘‘Blueprint’’ consortium FP7-282510), the European Research Council (EpiGePlas) and the SNF Sinergia program.

References

Aaboe, M., Birkenkamp-Demtroder, K., Wiuf, C., Sorensen, F. B., Thykjaer, T., Sauter, G., Jensen, K. M., Dyrskjot, L., and Orntoft, T. (2006). SOX4 expression in bladder carcinoma: clinical aspects and in vitro functional characterization. Cancer Res *66*, 3434-3442.

Acloque, H., Ocana, O. H., Matheu, A., Rizzoti, K., Wise, C., Lovell-Badge, R., and Nieto, M. A. (2011). Reciprocal repression between Sox3 and snail transcription factors defines embryonic territories at gastrulation. Dev Cell *21*, 546-558.

Ahn, S. G., Kim, H. S., Jeong, S. W., Kim, B. E., Rhim, H., Shim, J. Y., Kim, J. W., Lee, J. H., and Kim, I. K. (2002). Sox-4 is a positive regulator of Hep3B and HepG2 cells' apoptosis induced by prostaglandin (PG)A(2) and delta(12)-PGJ(2). Exp Mol Med *34*, 243-249.

Brabletz, T., Jung, A., Spaderna, S., Hlubek, F., and Kirchner, T. (2005). Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. Nat Rev Cancer *5*, 744-749.

Cao, Q., Yu, J., Dhanasekaran, S. M., Kim, J. H., Mani, R. S., Tomlins, S. A., Mehra, R., Laxman, B., Cao, X., Kleer, C. G.*, et al.* (2008). Repression of E-cadherin by the polycomb group protein EZH2 in cancer. Oncogene *27*, 7274-7284.

Chase, A., and Cross, N. C. (2011). Aberrations of EZH2 in cancer. Clin Cancer Res *17*, 2613-2618.

Chen, H., Tu, S. W., and Hsieh, J. T. (2005). Down-regulation of human DAB2IP gene expression mediated by polycomb Ezh2 complex and histone deacetylase in prostate cancer. J Biol Chem *280*, 22437-22444.

Cheung, M., Abu-Elmagd, M., Clevers, H., and Scotting, P. J. (2000). Roles of Sox4 in central nervous system development. Brain Res Mol Brain Res *79*, 180-191.

Chew, L. J., and Gallo, V. (2009). The Yin and Yang of Sox proteins: Activation and repression in development and disease. J Neurosci Res *87*, 3277-3287.

Christofori, G. (2006). New signals from the invasive front. Nature *441*, 444-450.

Crea, F., Hurt, E. M., Mathews, L. A., Cabarcas, S. M., Sun, L., Marquez, V. E., Danesi, R., and Farrar, W. L. (2011). Pharmacologic disruption of Polycomb Repressive Complex 2 inhibits tumorigenicity and tumor progression in prostate cancer. Mol Cancer *10*, 40.

Ezhkova, E., Pasolli, H. A., Parker, J. S., Stokes, N., Su, I. H., Hannon, G., Tarakhovsky, A., and Fuchs, E. (2009). Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. Cell *136*, 1122-1135.

Grunert, S., Jechlinger, M., and Beug, H. (2003). Diverse cellular and molecular mechanisms contribute to epithelial plasticity and metastasis. Nat Rev Mol Cell Biol *4*, 657-665.

Guo, W., Keckesova, Z., Donaher, J. L., Shibue, T., Tischler, V., Reinhardt, F., Itzkovitz, S., Noske, A., Zurrer-Hardi, U., Bell, G.*, et al.* (2012). Slug and Sox9 cooperatively determine the mammary stem cell state. Cell *148*, 1015-1028.

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646-674.

Herranz, N., Pasini, D., Diaz, V. M., Franci, C., Gutierrez, A., Dave, N., Escriva, M., Hernandez-Munoz, I., Di Croce, L., Helin, K.*, et al.* (2008). Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor. Mol Cell Biol *28*, 4772-4781.

Hong, C. S., and Saint-Jeannet, J. P. (2005). Sox proteins and neural crest development. Semin Cell Dev Biol *16*, 694-703.

Huber, M. A., Kraut, N., and Beug, H. (2005). Molecular requirements for epithelial-mesenchymal transition during tumor progression. Curr Opin Cell Biol *17*, 548-558.

Hur, W., Rhim, H., Jung, C. K., Kim, J. D., Bae, S. H., Jang, J. W., Yang, J. M., Oh, S. T., Kim, D. G., Wang, H. J.*, et al.* (2010). SOX4 overexpression regulates the p53-mediated apoptosis in hepatocellular carcinoma: clinical implication and functional analysis in vitro. Carcinogenesis *31*, 1298-1307.

Ikushima, H., Todo, T., Ino, Y., Takahashi, M., Miyazawa, K., and Miyazono, K. (2009). Autocrine TGF-beta signaling maintains tumorigenicity of glioma-initiating cells through Sry-related HMG-box factors. Cell Stem Cell *5*, 504-514.

Kalluri, R. (2009). EMT: when epithelial cells decide to become mesenchymal-like cells. J Clin Invest *119*, 1417-1419.

Kalluri, R., and Weinberg, R. A. (2009). The basics of epithelial-mesenchymal transition. J Clin Invest *119*, 1420-1428.

Kang, Y., and Massague, J. (2004). Epithelial-mesenchymal transitions: twist in development and metastasis. Cell *118*, 277-279.

Kleer, C. G., Cao, Q., Varambally, S., Shen, R., Ota, I., Tomlins, S. A., Ghosh, D., Sewalt, R. G., Otte, A. P., Hayes, D. F.*, et al.* (2003). EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. Proc Natl Acad Sci U S A *100*, 11606-11611.

Kloeker, S., Major, M. B., Calderwood, D. A., Ginsberg, M. H., Jones, D. A., and Beckerle, M. C. (2004). The Kindler syndrome protein is regulated by transforming growth factor-beta and involved in integrin-mediated adhesion. The Journal of biological chemistry *279*, 6824-6833.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lehembre, F., Yilmaz, M., Wicki, A., Schomber, T., Strittmatter, K., Ziegler, D., Kren, A., Went, P., Derksen, P. W., Berns, A.*, et al.* (2008). NCAM-induced focal adhesion assembly: a functional switch upon loss of E-cadherin. EMBO J *27*, 2603-2615.

Liao, Y. L., Sun, Y. M., Chau, G. Y., Chau, Y. P., Lai, T. C., Wang, J. L., Horng, J. T., Hsiao, M., and Tsou, A. P. (2008). Identification of SOX4 target genes using phylogenetic footprinting-based prediction from expression microarrays suggests that overexpression of SOX4 potentiates metastasis in hepatocellular carcinoma. Oncogene.

Lioubinski, O., Muller, M., Wegner, M., and Sander, M. (2003). Expression of Sox transcription factors in the developing mouse pancreas. Dev Dyn *227*, 402-408.

Liu, P., Ramachandran, S., Ali Seyed, M., Scharer, C. D., Laycock, N., Dalton, W. B., Williams, H., Karanam, S., Datta, M. W., Jaye, D. L., and Moreno, C. S. (2006). Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells. Cancer Res *66*, 4011-4019.

Maeda, M., Johnson, K. R., and Wheelock, M. J. (2005). Cadherin switching: essential for behavioral but not morphological changes during an epithelium-to-mesenchyme transition. Journal of cell science *118*, 873-887.

Mani, S. A., Guo, W., Liao, M. J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., Brooks, M., Reinhard, F., Zhang, C. C., Shipitsin, M.*, et al.* (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell *133*, 704-715.

Mills, A. A. (2010). Throwing the cancer switch: reciprocal roles of polycomb and trithorax proteins. Nature reviews Cancer *10*, 669-682.

Min, J., Zaslavsky, A., Fedele, G., McLaughlin, S. K., Reczek, E. E., De Raedt, T., Guney, I., Strochlic, D. E., Macconaill, L. E., Beroukhim, R.*, et al.* (2010). An oncogene-tumor suppressor cascade drives metastatic prostate cancer by coordinately activating Ras and nuclear factor-kappaB. Nat Med *16*, 286-294.

Minn, A. J., Gupta, G. P., Siegel, P. M., Bos, P. D., Shu, W., Giri, D. D., Viale, A., Olshen, A. B., Gerald, W. L., and Massague, J. (2005). Genes that mediate breast cancer metastasis to lung. Nature *436*, 518-524.

Morel, A. P., Lievre, M., Thomas, C., Hinkal, G., Ansieau, S., and Puisieux, A. (2008). Generation of breast cancer stem cells through epithelial-mesenchymal transition. PLoS One *3*, e2888.

Nieto, M. A. (2010). The Ins and Outs of the Epithelial to Mesenchymal Transition in Health and Disease. Annu Rev Cell Dev Biol.

Penzo-Mendez, A. I. (2010). Critical roles for SoxC transcription factors in development and cancer. Int J Biochem Cell Biol *42*, 425-428.

Polyak, K., and Weinberg, R. A. (2009). Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nat Rev Cancer *9*, 265-273.

Pramoonjago, P., Baras, A. S., and Moskaluk, C. A. (2006). Knockdown of Sox4 expression by RNAi induces apoptosis in ACC3 cells. Oncogene *25*, 5626-5639.

Restivo, A., Piacentini, G., Placidi, S., Saffirio, C., and Marino, B. (2006). Cardiac outflow tract: a review of some embryogenetic aspects of the conotruncal region of the heart. Anat Rec A Discov Mol Cell Evol Biol *288*, 936-943.

Rhodes, D. R., Sanda, M. G., Otte, A. P., Chinnaiyan, A. M., and Rubin, M. A. (2003). Multiplex biomarker approach for determining risk of prostate-specific antigen-defined recurrence of prostate cancer. J Natl Cancer Inst *95*, 661-668.

Sauvageau, M., and Sauvageau, G. (2010). Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. Cell Stem Cell *7*, 299-313.

Scharer, C. D., McCabe, C. D., Ali-Seyed, M., Berger, M. F., Bulyk, M. L., and Moreno, C. S. (2009). Genome-wide promoter analysis of the SOX4 transcriptional network in prostate cancer cells. Cancer Res *69*, 709-717.

Schilham, M. W., Moerer, P., Cumano, A., and Clevers, H. C. (1997). Sox-4 facilitates thymocyte differentiation. Eur J Immunol *27*, 1292-1295.

Schilham, M. W., Oosterwegel, M. A., Moerer, P., Ya, J., de Boer, P. A., van de Wetering, M., Verbeek, S., Lamers, W. H., Kruisbeek, A. M., Cumano, A., and Clevers, H. (1996). Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. Nature *380*, 711-714.

Schmidt, M., Bohm, D., von Torne, C., Steiner, E., Puhl, A., Pilch, H., Lehr, H. A., Hengstler, J. G., Kolbl, H., and Gehrmann, M. (2008a). The humoral immune system has a key prognostic impact in node-negative breast cancer. Cancer Res *68*, 5405-5413.

Shen, R., Pan, S., Qi, S., Lin, X., and Cheng, S. (2010). Epigenetic repression of microRNA-129-2 leads to overexpression of SOX4 in gastric cancer. Biochem Biophys Res Commun *394*, 1047-1052.

Shin, M. S., Fredrickson, T. N., Hartley, J. W., Suzuki, T., Akagi, K., and Morse, H. C., 3rd (2004). High-throughput retroviral tagging for identification of genes involved in initiation and progression of mouse splenic marginal zone lymphomas. Cancer Res *64*, 4419-4427.

Sinner, D., Kordich, J. J., Spence, J. R., Opoka, R., Rankin, S., Lin, S. C., Jonatan, D., Zorn, A. M., and Wells, J. M. (2007). Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells. Mol Cell Biol *27*, 7802-7815.

Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. Nature reviews Cancer *6*, 846-856.

Su, I. H., Basavaraj, A., Krutchinsky, A. N., Hobert, O., Ullrich, A., Chait, B. T., and Tarakhovsky, A. (2003). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat Immunol *4*, 124-131.

Suzuki, H., Forrest, A.R., van Nimwegen, E., Daub, C.O., Balwierz, P.J., Irvine, K.M., Lassmann, T., Ravasi, T., Hasegawa, Y., de Hoon, M.J., *et al.* (2009). The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nat Genet *41*, 553-562.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tavazoie, S. F., Alarcon, C., Oskarsson, T., Padua, D., Wang, Q., Bos, P. D., Gerald, W. L., and Massague, J. (2008). Endogenous human microRNAs that suppress breast cancer metastasis. Nature *451*, 147-152.

Thiery, J. P., and Morgan, M. (2004). Breast cancer progression with a Twist. Nat Med *10*, 777-778.

Thiery, J. P., and Sleeman, J. P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. Nat Rev Mol Cell Biol *7*, 131-142.

van de Wetering, M., Oosterwegel, M., van Norren, K., and Clevers, H. (1993). Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. EMBO J *12*, 3847-3854.

Varambally, S., Cao, Q., Mani, R. S., Shankar, S., Wang, X., Ateeq, B., Laxman, B., Cao, X., Jing, X., Ramnarayanan, K.*, et al.* (2008). Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. Science *322*, 1695-1699.

Waldmeier, L., Meyer-Schaller, N., Diepenbruck, M., and Christofori, G. (2012). Py2T murine breast cancer cells, a versatile model of TGFbeta-induced EMT in vitro and in vivo. PLoS ONE *7*, e48651.

Wilson, M. E., Yang, K. Y., Kalousova, A., Lau, J., Kosaka, Y., Lynn, F. C., Wang, J., Mrejen, C., Episkopou, V., Clevers, H. C., and German, M. S. (2005). The HMG box transcription factor Sox4 contributes to the development of the endocrine pancreas. Diabetes *54*, 3402-3409.

Yang, X., Pursell, B., Lu, S., Chang, T. K., and Mercurio, A. M. (2009). Regulation of beta 4-integrin expression by epigenetic modifications in the mammary gland and during the epithelial-to-mesenchymal transition. Journal of cell science *122*, 2473-2480.

Zhang, J., Liang, Q., Lei, Y., Yao, M., Li, L., Gao, X., Feng, J., Zhang, Y., Gao, H., Liu, D. X.*, et al.* (2012). SOX4 induces epithelial-mesenchymal transition and contributes to breast cancer progression. Cancer research *72*, 4597-4608.

Figure Legends

Figure 1. Sox4 is required for EMT.

(**A**) Sox4 mRNA levels were determined by quantitative RT-PCR in NMuMG cells transfected with either control siRNA (siControl) or with siRNA against Sox4 (siSox4) in the absence or presence of TGFβ for 2 days. Fold changes are shown in comparison to cells transfected with control siRNA in the absence of TGFβ.

(**B**) NMuMG cells transfected with either siControl or with siSox4 were treated with TGFβ for the days indicated and their morphology was evaluated by phase contrast microscopy. Size bar, 100μm.

(**C**) Immunoflorescence microscopy analysis of changes in the localization and expression levels of marker proteins during EMT. NMuMG cells transfected with either siControl or with siSox4 were left untreated or treated with TGFβ for 2 days and stained with antibodies against the epithelial markers E-cadherin and ZO-1, against the mesenchymal marker N-cadherin, against paxillin to detect focal adhesion plaques, and with phalloidin to visualize the actin cytoskeleton. Size bar, 50μm.

(**D**) Expression of E-cadherin, N-cadherin and fibronectin was determined by immunoblotting during TGFβ-induced EMT in NMuMG cells transfected either with siSox4 or siControl. Immunoblotting for actin was used as a loading control.

(**E**) Morphology of long-term TGFβ-treated NMuMG (NMuMG-LT) cells transfected with either siControl or siSox4 as evaluated by phase contrast microscopy. Size bar, 100μm.

Statistical values were calculated using a paired, two-tailed t-test. ** = $p \leq 0.01$; *** = $p \leq 0.001$; error bars = ±SD. See also Figure S1.

Figure 2. Sox4 is required for cell survival and cell migration during TGFβ-induced EMT.

(**A, B**) siRNA or shRNA-mediated ablation of Sox4 expression (siSox4 and shSox4, respectively) during TGFβ treatment of NMuMG (A) and Py2T (B) cells results in a significant decrease in cell numbers as compared to cells transfected with control siRNA or shRNA (siControl and shControl, respectively).

(**C, D**) NMuMG cells (C) or Py2T cells (D) transfected with either siControl and shControl or with siSox4 and shSox4 were treated with TGFβ for the days indicated (d0 to d4). Cells were stained with propidium iodide (PI), and the percentages of cells in G0/G1 and S-G2/M phases of the cell cycle were determined by flow cytometry.

(**E**) NMuMG cells transfected with either siControl or with siSox4 were treated with TGFβ for the days indicated, and the rates of apoptosis were determined by Annexin-V staining and flow cytometry.

(**F**) shSox4-expressing and shControl-expressing Py2T murine breast cancer cells were treated for 15 days with TGFβ and cell migration was determined in a modified Boyden chamber migration assay 20 hours after seeding and using 20% FBS as a chemoattractant.

Statistical values were calculated using an unpaired, two-tailed t-test. * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$; error bars = ±SD.

Figure 3. Ectopic expression of Sox4 promotes EMT and cell migration.

(A) NMuMG cells were transiently transfected with pcDNA3-HA and pcDNA3-HA-Sox4 constructs and cellular morphology was evaluated. Note the formation of filopodia and cell dispersal upon Sox4 expression (HA-Sox4) as compared to empty vector control (Empty). Size bar, 100μm.

(B) The expression of E-cadherin (Ecad), N-cadherin (Ncad), fibronectin (Fn1), Ezh2 and Sox4 was determined by quantitative RT-PCR in NMuMG cells transiently transfected with HA-Sox4 or with empty vector (Empty) in the absence of TGFβ. Fold changes are shown in comparison to cells transfected with empty vector.

(C) The expression of E-cadherin, N-cadherin, fibronectin, Ezh2 and HA-tagged Sox4 was determined by immunoblotting of NMuMG cells transiently transfected with empty vector or with HA-Sox4. Actin was used as loading control.

(D) The localization of E-cadherin was determined by immunofluorescence staining in NMuMG cells transiently transfected with empty vector control or with HA-Sox4. The cells marked with an arrow express Sox4 (HA) and have lost E-cadherin at their cell membranes. Size bar, 50μm.

(**E**) Cell scattering and mesenchymal morphology is observed even in the absence of TGFβ in Py2T cell clones stably expressing HA-Sox4 (Clone 3 and 5) as compared to empty vector-transfected control cells (Empty). Size bar, 100μm.

(**F**) The expression of E-cadherin (Ecad), N-cadherin (Ncad), fibronectin (Fn1), and Sox4 was determined by quantitative RT-PCR in Py2T cell clones either stably transfected with HA-Sox4 (Clone 3 and 5) or with control vector (Empty) in the absence of TGFβ. Fold changes are shown in comparison to cells transfected with the control vector.

(**G**) Immunoblotting analysis of E-cadherin, N-cadherin, fibronectin, Ezh2 and HA-Sox4 (HA) during TGFβ-induced EMT in Py2T cell clones either stably transfected with HA-Sox4 (Clone 3

and 5) or with empty control vector (Empty) in the absence of TGFβ. Vinculin was used as loading control.

(**H**) Cell migration of Py2T cell clones stably transfected with HA-Sox4 (Clone 3 and 5) or transfected with control vector (Empty) was analyzed in a trans-well migration assay in the absence of TGFβ after 20 hours of cell seeding with 20% FBS as a chemoattractant.

Statistical values were calculated using an unpaired, two-tailed t-test. * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$; error bars = ±SD.

Figure 4. Sox4 is critical for breast carcinogenesis and metastasis.

(**A**) Correlation of the levels of Sox4 expression and prognosis in the "Schmidt" array data set of lymph node-negative breast cancers (Schmidt et al., 2008). Shown is a Kaplan-Meier survival curve for patient samples classified for having high Sox4 expression and low Sox4 expression to assess metastasis-free survival.

(**B**) shSox4-expressing and shControl-expressing Py2T murine breast cancer cells were injected into the mammary fat pad of nude mice and the resulting tumors were analyzed by immunohistochemistry for Sox4 expression. High magnification as indicated by boxed areas shows that shControl-transfected cells localize Sox4 primarily in the nucleus and exhibit spindle-like cell morphology, while shSox4-expressing cells lack significant Sox4 expression and exhibit a more differentiated morphology.

(**C**) shSox4-expressing and shControl-expressing Py2T cells were injected into the mammary fat pad of nude mice and tumor growth in individual mice was measured over time (top panel). Tumor weight was assessed after sacrificing the tumor-bearing mice 23 days after injection (bottom panel).

(**D**) Metastatic spread of firefly luciferase-expressing Py2T cells from the primary tumors analyzed in panel C was determined by measuring luciferase activity in extracts of lymph nodes (LN), lungs and livers of transplanted mice. Luciferase activity levels in the various organs were divided by the primary tumor weights within the same mice to establish the metastatic index. Numbers indicate the fold-differences between Sox4-expressing (shControl) and Sox4-depleted (shSox4) Py2T cells. Error bars, ±SE.

(**E**) Primary tumor growth of shControl and shSox4 Py2T cells upon subcutaneous injection into nude mice (left panel). Tumor weight was assessed 27 days after implantation (right panel).

Statistical values were calculated using an unpaired, two-tailed t-test. * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$; error bars = ±SEM. See also Figure S2.

Figure 5. Sox4 directly regulates *Ezh2* gene expression.

(**A-B**) NMuMG cells (A) and Py2T cells (B) were transiently transfected with a construct encoding HA-Sox4 or empty vector and used for Chromatin Immunoprecipitation (ChIP) in the absence and presence of TGFβ for 2 days using HA-specific antibodies. Precipitated DNA fragments corresponding to *Ezh2* promoter sequences were detected by quantitative PCR to assess Sox4 occupancy at this region. Fold changes are shown in comparison to a PCR reaction with primers specific for an intergenic region.

(**C**) Levels of Ezh2 mRNA were determined by quantitative RT-PCR in NMuMG cells or Py2T cells transfected with control siRNA (siControl) or siRNA against Sox4 (siSox4) in the presence of TGFβ for 2 days. Fold changes are shown in comparison to siControl-transfected cells (top panels). Global H3K27me3 levels were determined by immunoblotting in NMuMG cells or Py2T cells transfected with siControl or siSox4 and treated with TGFβ for 2 days. Total histone H3 was used as loading control (bottom panels).

(**D**) Sox4 regulates Ezh2 promoter activity. NMuMG cells were transiently transfected with siControl or siSox4 and with a pGL3 luciferase reporter plasmid containing a 1.9kb fragment of the Ezh2 promoter controlling the expression of *firefly* luciferase and then treated with or without TGFβ for 2 days. *Firefly* luciferase activity was normalized to co-transfected *renilla* luciferase activity (relative luminescence).

(E, F) Sox4 and Ezh2 expression levels were determined by quantitative RT-PCR after transient transfection of NMuMG (E) and Py2T cells (F) with siControl or siSox4 during the first hours of TGFβ treatment (0, 2, 4 and 8 hours).

Statistical values were calculated using a paired, two-tailed t-test. * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$; error bars = ±SD. See also Figure S3 and Tables S1 and S2.

Figure 6. Depletion of Ezh2 phenocopies the ablation of Sox4 during TGFβ-induced EMT.

(**A**) NMuMG cells transfected with either control siRNA (siControl) or with siRNAs against Ezh2 or Ezh1 and Ezh2 together (siEzh1/2) were treated with TGFβ for the days indicated and their morphology was evaluated by phase contrast microscopy. Size bar, 100μm.

(**B**) Expression of E-cadherin, N-cadherin, ZO-1 and Ezh2 was determined by immunoblotting analysis during TGFβ-induced EMT in NMuMG cells transfected either with siEzh2 or siControl. Immunoblotting for actin was used as a loading control.

(**C**) Immunoflorescence microscopy analysis of NMuMG cells transfected with either siControl or with siEzh2 or siEzh1/2. Cells were treated with TGFβ for 7 days and stained with antibodies

against E-cadherin, ZO-1, and fibronectin, and with phalloidin to visualize the actin cytoskeleton. Size bar, 50μm.

(**D**) NMuMG cells transfected with either siControl or with siEzh2 were treated with TGFβ for the days indicated, and the rates of apoptosis were determined by Annexin-V staining and flow cytometry.

(**E**) siControl and siEzh2-transfected NMuMG cells were treated with TGFβ for the days indicated. Cells were stained with propidium iodide (PI), and the percentages of cells in G0/G1 and S-G2/M phases of the cell cycle were determined by flow cytometry.

(**F**) siRNA-mediated ablation of Ezh2 expression during TGFβ treatment of NMuMG cells results in a significant decrease in cell numbers as compared to cells transfected with control siRNA.

(G) Kaplan-Meier survival curve for patient samples of the "Schmidt" array data set of lymph node-negative breast cancers (Schmidt et al., 2008) classified for having concomitant low Sox4 and Ezh2 expression and high Sox4 and Ezh2 expression to assess metastasis free survival.

Statistical values were calculated using an unpaired, two-tailed t-test. * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$; error bars = ±SD. See also Figure S4

Figure 7. Forced expression of Ezh2 overcomes the lack of Sox4 and restores TGFβ-induced EMT.

(**A**) While ablation of Sox4 expression (siSox4) repressed the TGFβ-induced mesenchymal morphology of NMuMG cells observed in siControl-transfected cells, transient expression of Ezh2 in Sox4-ablated cells (siSox4 + Ezh2) restored EMT and mesenchymal cell morphology. Cells were treated with TGFβ for the indicated times. Size bar, 100μm.

(**B, C**) mRNA expression levels of E-cadherin (C) and fibronectin (D) were determined after Ezh2 overexpression (+Ezh2) in control (shControl) and Sox4-ablated (shSox4) Py2T cells.

Statistical values were calculated using an unpaired, two-tailed t-test. * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$. See also Figure S5.

Figure 8. TGFβ-induced EMT accompanies genome-wide reprogramming of the Polycomb-associated mark H3K27me3.

(**A**) TGFβ-induced EMT accompanies genome-wide reprogramming of the Polycomb-associated mark H3K27me3. Genome-wide H3K27me3 enrichment above background was calculated for 2kb-wide windows overlapping by 1kb. To extract those H3K27me3-regions that are most relevant for EMT dynamics, for each region both the variance in H3K27me3 levels during the

progressive stages of EMT (horizontal axis) and the reproducibility of the H3K27me3 dynamics across biological replicates (vertical) axis was calculated and 970 genes with high variance and high reproducibility were selected. Comparison of these regions with RefSeq transcripts show that they fall within 1kb upstream and downstream of 301 distinct target genes.

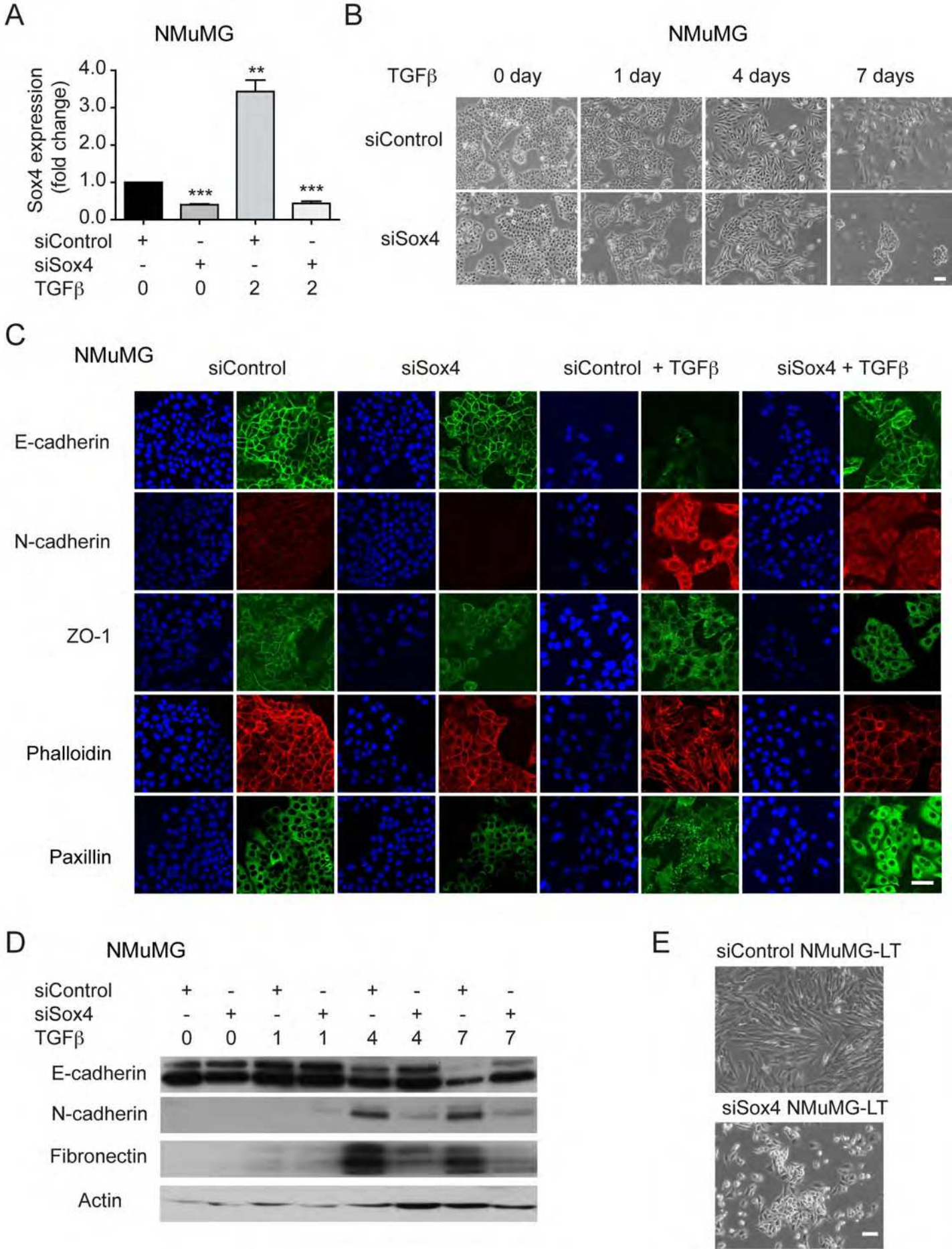
(B, C) ChIP was performed using an H3K27me3-specific antibody in NMuMG cells and the enrichment of H3K27me3 was determined by quantitative PCR at the promoters of the *Mcam*, *Pdgfrb* and *Sgsh* genes during TGFβ-induced EMT (B) and at the promoters of the *Itga5*, *Col1a1* and *St6galnac4* genes in the absence of TGFβ (C). Average enrichments from independent assays are plotted on the *y* axis as the ratio of precipitated DNA relative to the total input DNA and further normalized to the *Gapdh* house-keeping gene promoter.

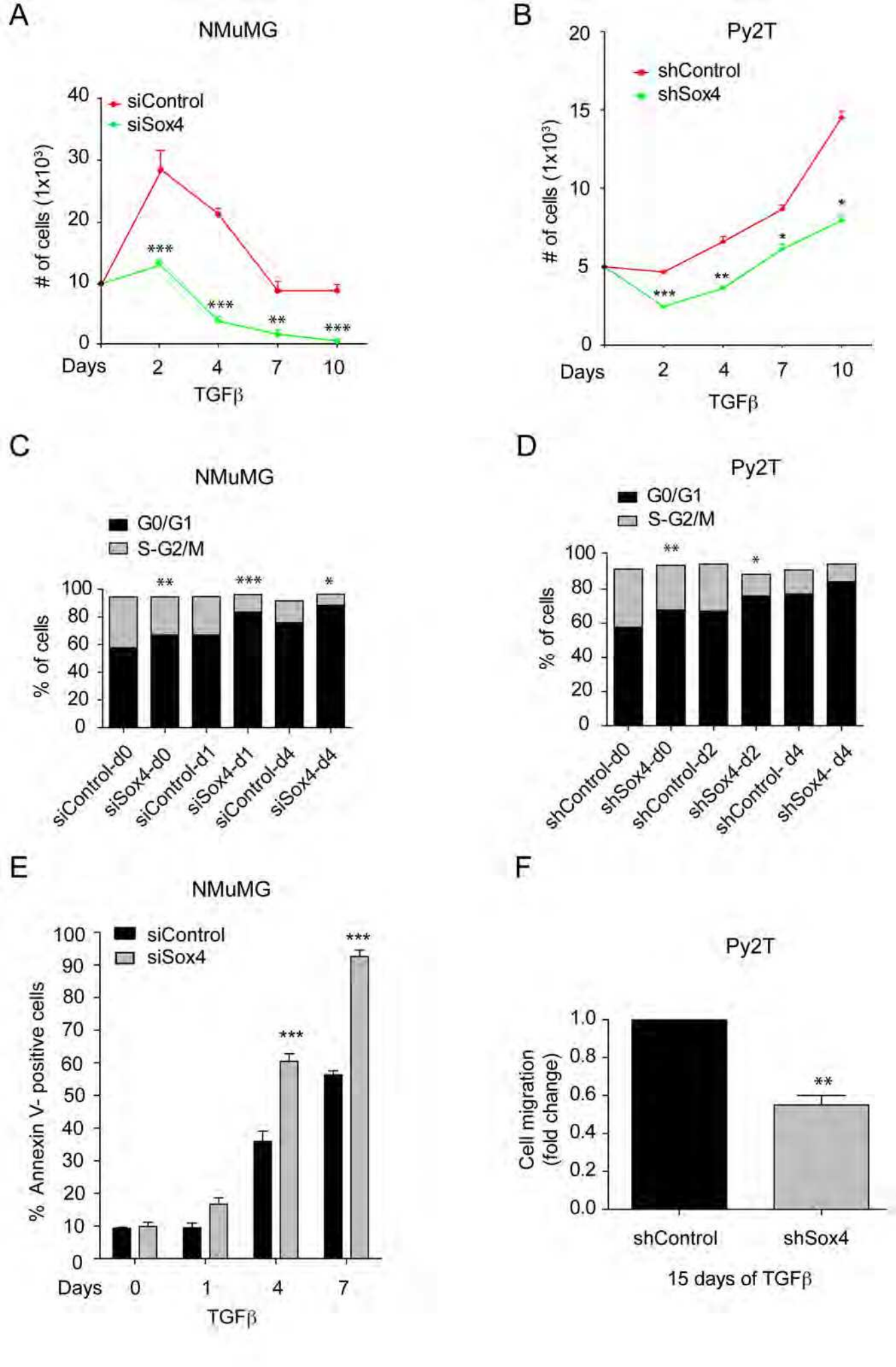
(**D**) ChIP was performed using an H3K27me3-specific antibody in NMuMG cells transiently transfected with control siRNA (siControl) or siRNA against Ezh1/2 (siEzh1/2), and quantitative PCR was performed using promoter-specific primers as indicated to determine enrichment of H3K27me3 at the *Mcam*, *Pdgfrb*, and *Itga5* promoters and normalized to the *Gapdh* house-keeping gene promoter.

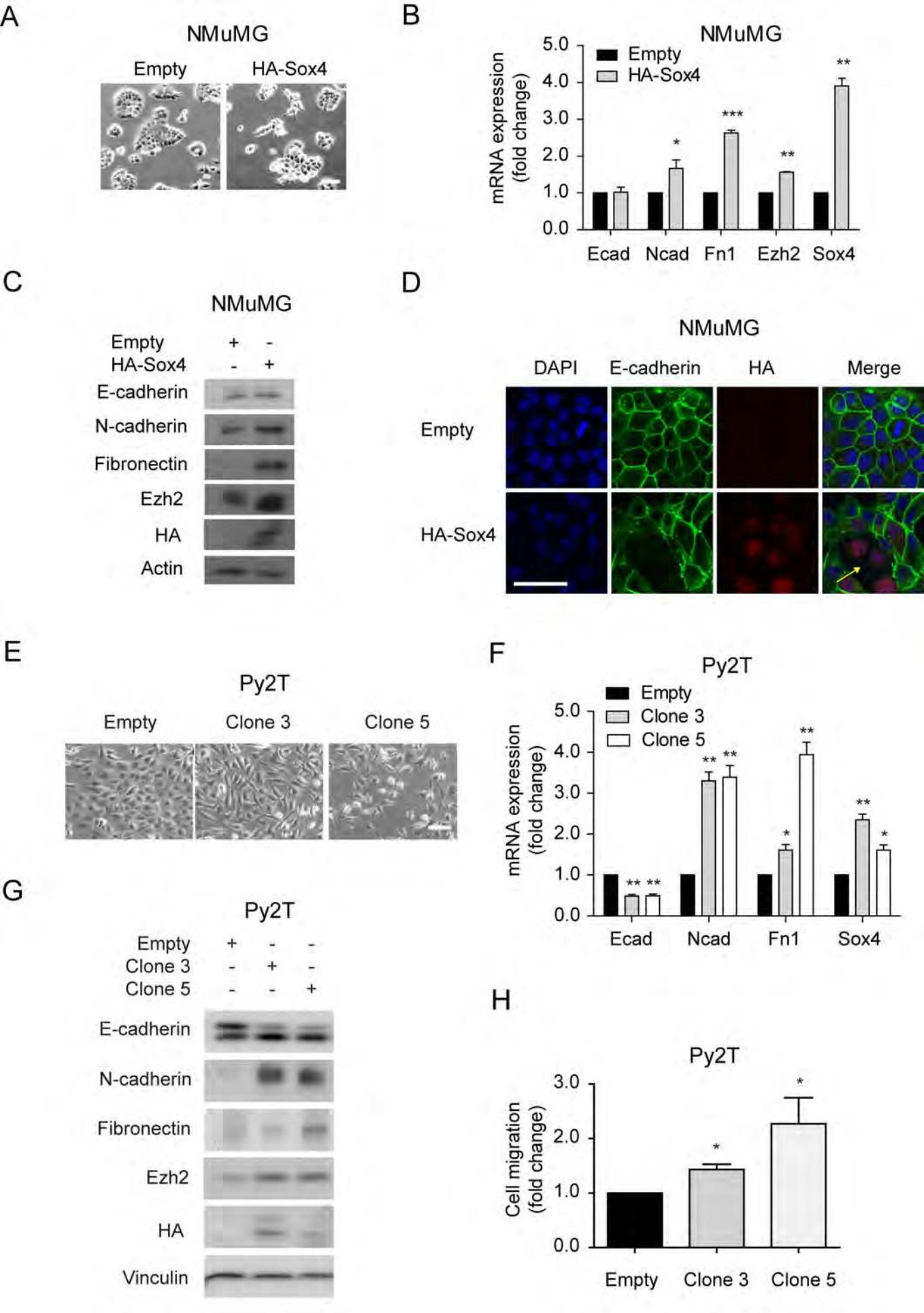
(**E - G**) NMuMG cells transiently transfected with control siRNA (siControl) or siRNA against Ezh1/2 (siEzh1/2) were analyzed by quantitative RT-PCR analysis for the mRNA expression levels of *Mcam* (E) and *Pdgfrb* (F) which already carry high H3K27me3 marks in untreated NMuMG cells and lose this mark during EMT, and of *Sgsh* (G) which gains this mark upon TGFβ treatment for 1, 4, 7 and 10 days.

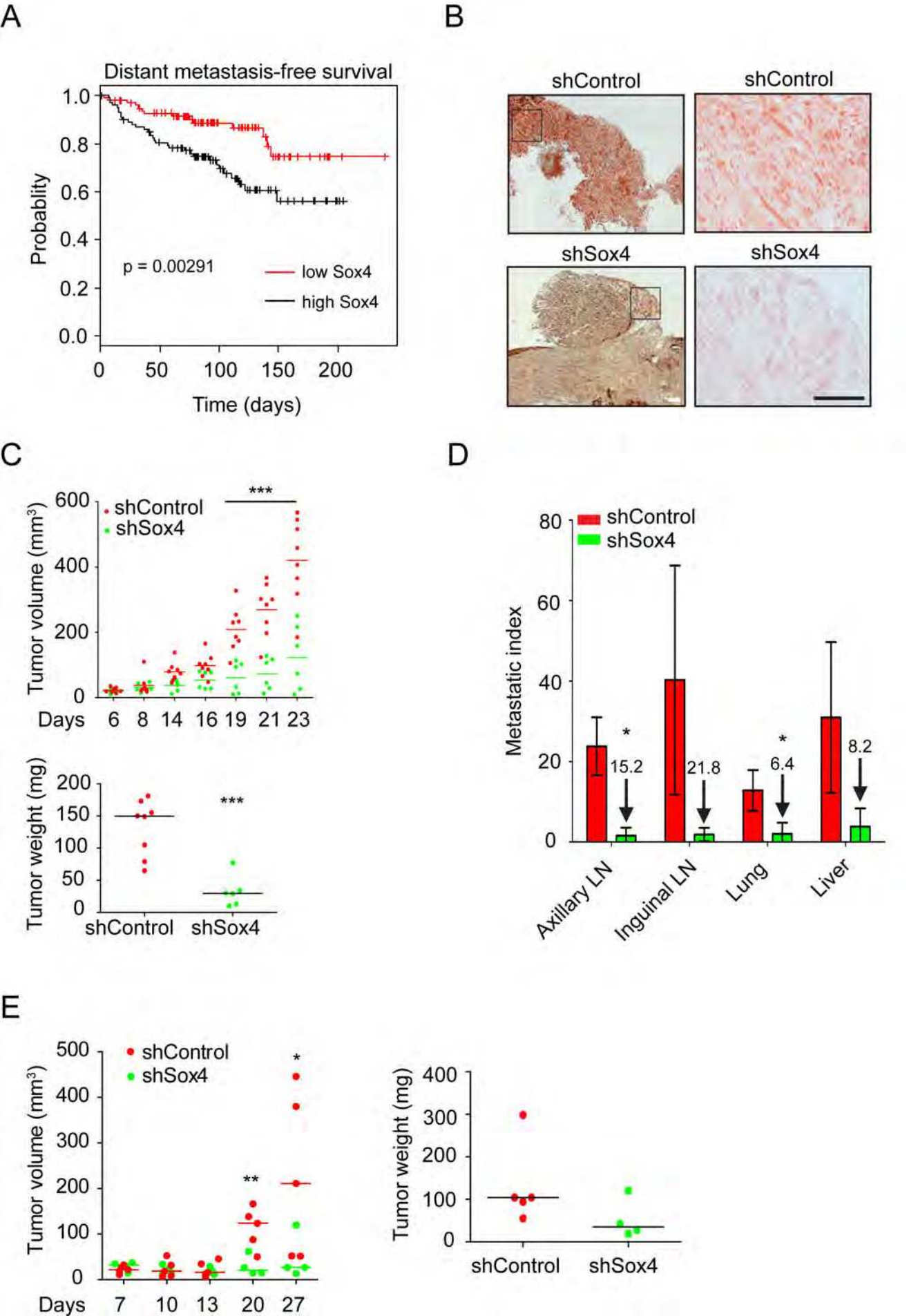
(**H**) Validation of selected genes enriched in H3K27me3 marks and their transcriptional de-repression after Ezh1/2 ablation in NMuMG cells in the absence of TGFβ.

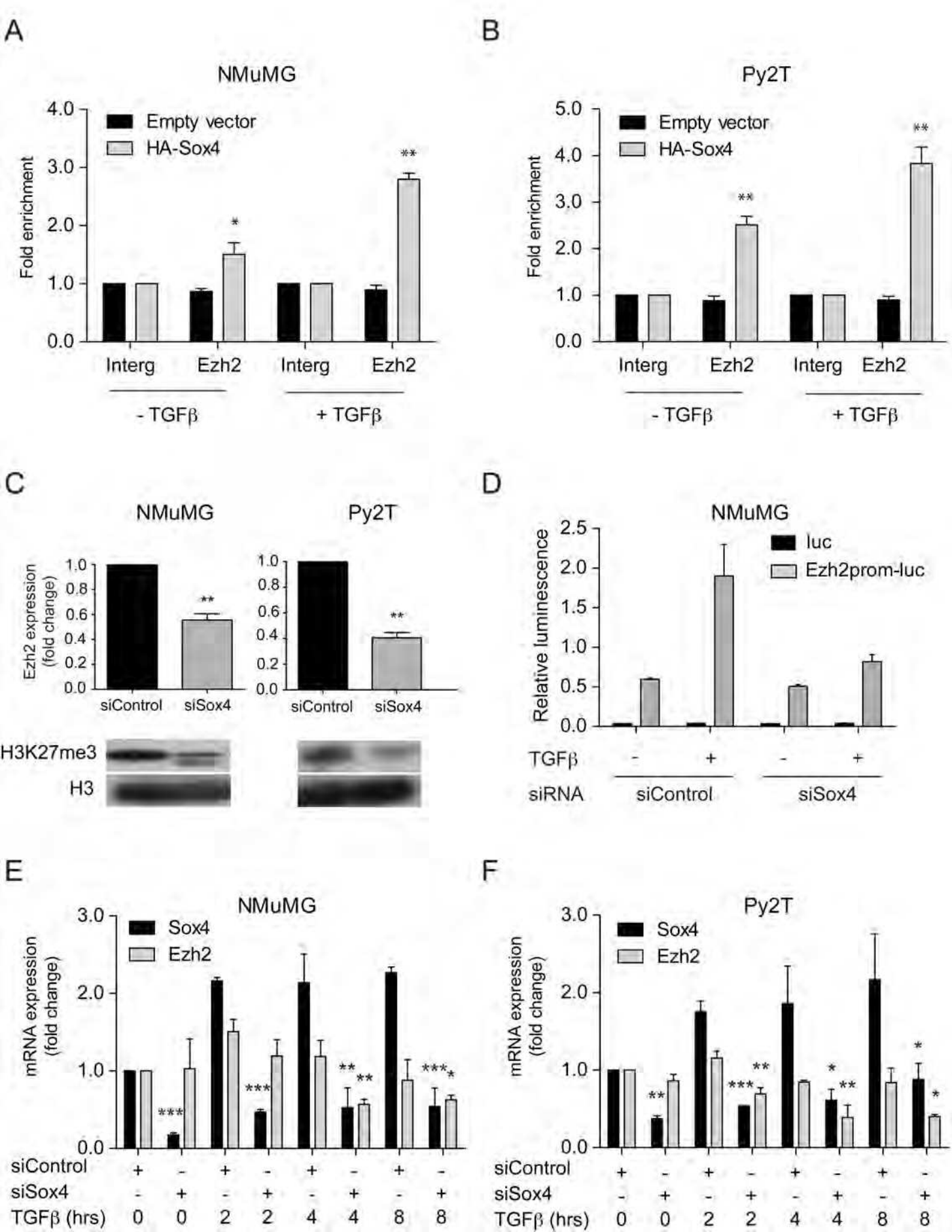
Statistical values were calculated using a paired, two-tailed t-test. * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$; error bars = ±SD. See also Figure S6 and Tables S1 and S2.

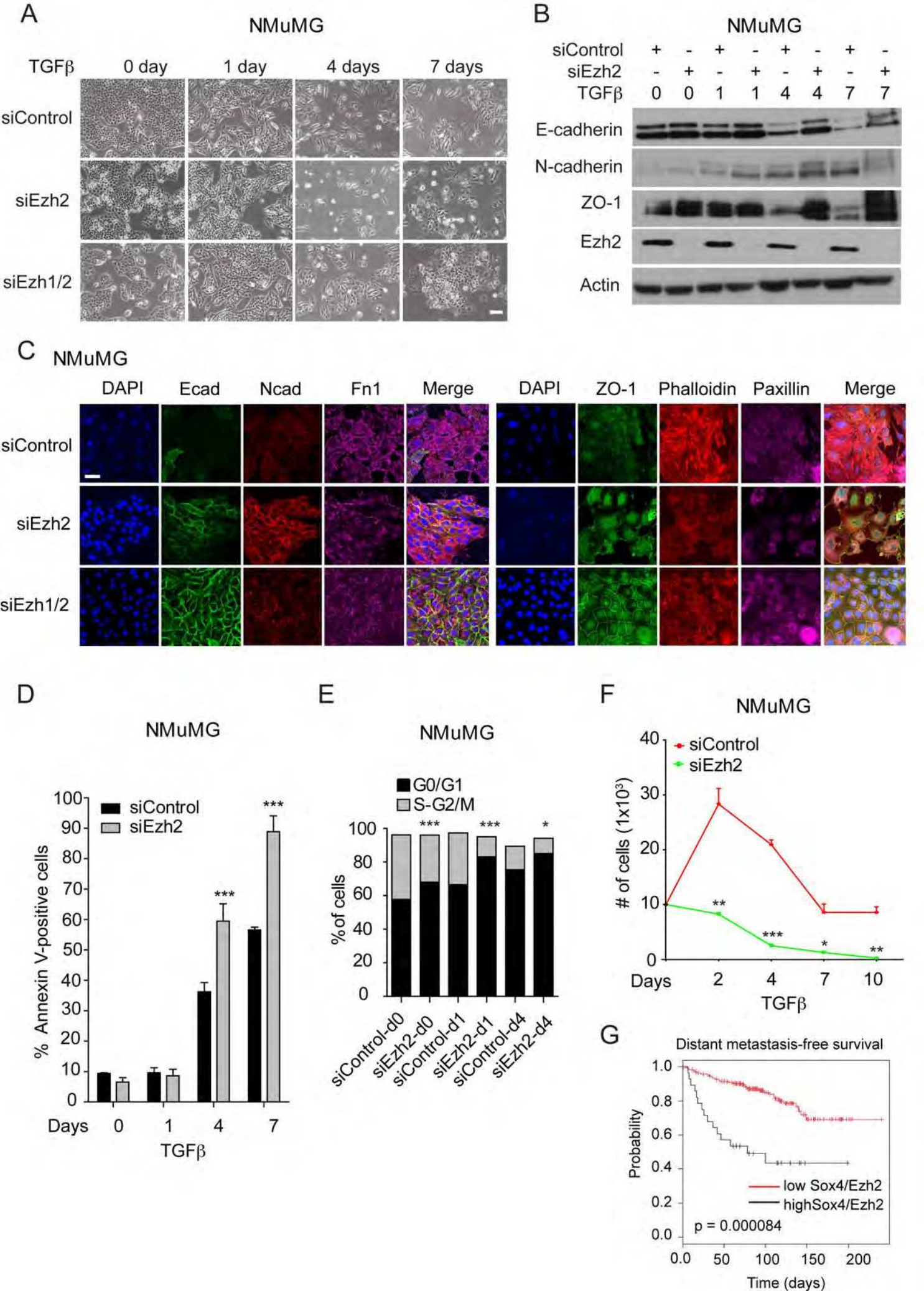

A
NMuMG
Sox4 expression (fold change)
4.0
3.0
2.0
1.0
0.0
**
***
***
siControl + - + -
siSox4 - + - +
TGFβ 0 0 2 2
B
NMuMG
TGFβ
0 day
1 day
4 days
7 days
siControl
siSox4
C
NMuMG
siControl
siSox4
siControl + TGFβ
siSox4 + TGFβ
E-cadherin
N-cadherin
ZO-1
Phalloidin
Paxillin
D
NMuMG
siControl + - + - + - + -
siSox4 - + - + - + - +
TGFβ 0 0 1 1 4 4 7 7
E-cadherin
N-cadherin
Fibronectin
Actin
E
siControl NMuMG-LT
siSox4 NMuMG-LT

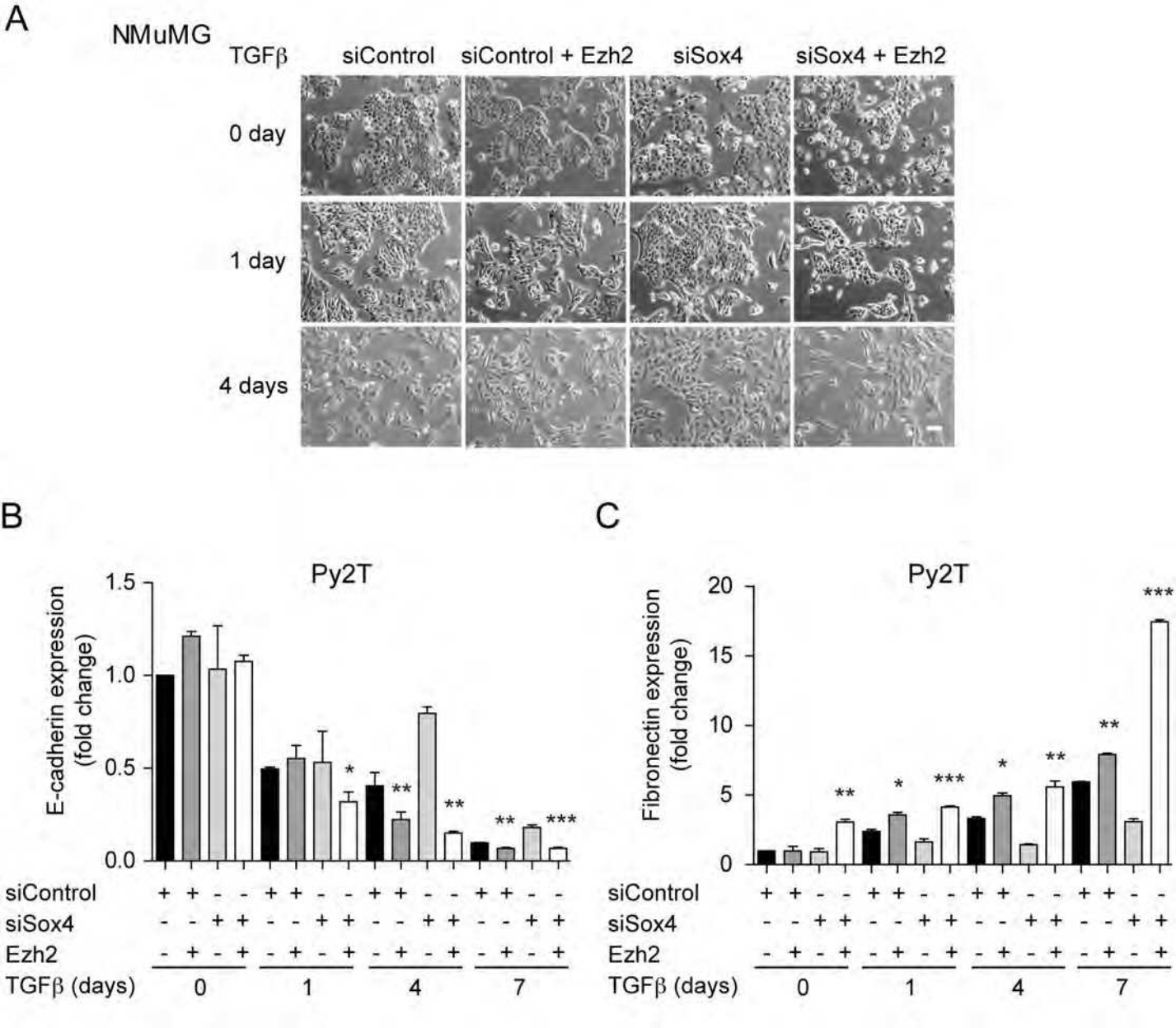

A
NMuMG
siControl
siSox4
# of cells (1x10³)
Days
TGFβ
B
Py2T
shControl
shSox4
# of cells (1x10³)
Days
TGFβ
C
NMuMG
G0/G1
S-G2/M
% of cells
siControl-d0
siSox4-d0
siControl-d1
siSox4-d1
siControl-d4
siSox4-d4
D
Py2T
G0/G1
S-G2/M
% of cells
shControl-d0
shSox4-d0
shControl-d2
shSox4-d2
shControl- d4
shSox4- d4
E
NMuMG
siControl
siSox4
% Annexin V- positive cells
Days
TGFβ
F
Py2T
Cell migration (fold change)
shControl
shSox4
15 days of TGFβ

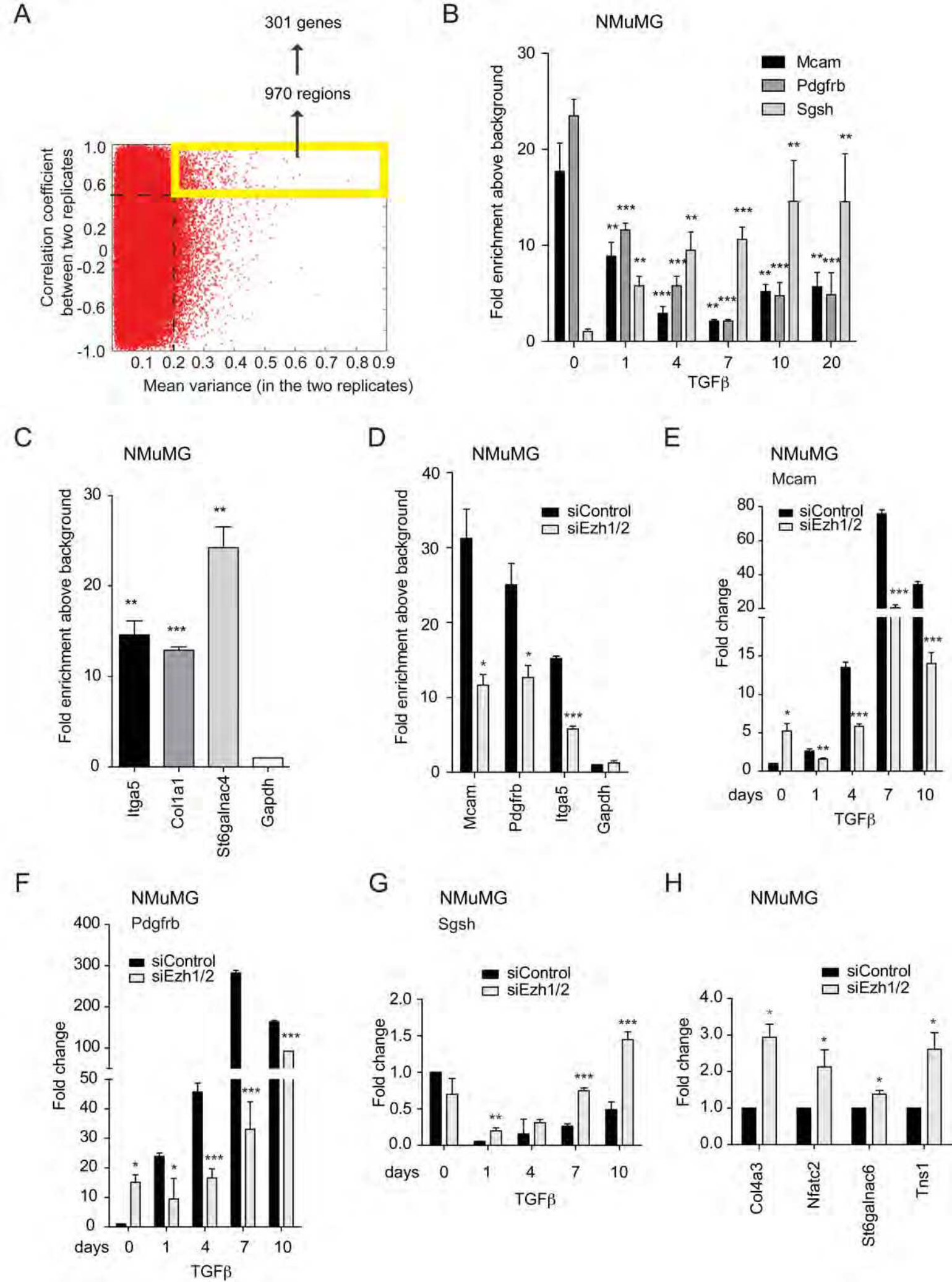

A
NMuMG
Empty
HA-Sox4
B
NMuMG
Empty
HA-Sox4
mRNA expression (fold change)
5.0
4.0
3.0
2.0
1.0
0.0
Ecad
Ncad
Fn1
Ezh2
Sox4
C
NMuMG
Empty + -
HA-Sox4 - +
E-cadherin
N-cadherin
Fibronectin
Ezh2
HA
Actin
D
NMuMG
DAPI
E-cadherin
HA
Merge
Empty
HA-Sox4
E
Py2T
Empty
Clone 3
Clone 5
F
Py2T
Empty
Clone 3
Clone 5
mRNA expression (fold change)
Ecad
Ncad
Fn1
Sox4
G
Py2T
Empty + - -
Clone 3 - + -
Clone 5 - - +
E-cadherin
N-cadherin
Fibronectin
Ezh2
HA
Vinculin
H
Py2T
Cell migration (fold change)
3.0
2.0
1.0
0.0
Empty
Clone 3
Clone 5


A
Distant metastasis-free survival
Probablity
1.0
0.8
0.6
0.4
0.2
0.0
p = 0.00291
low Sox4
high Sox4
0
50
100
150
200
Time (days)
B
shControl
shControl
shSox4
shSox4
C
Tumor volume (mm3)
600
400
200
0
shControl
shSox4
***
Days
6
8
14
16
19
21
23
Tumor weight (mg)
200
150
100
50
0
***
shControl
shSox4
D
Metastatic index
80
60
40
20
0
shControl
shSox4
*
15.2
21.8
*
6.4
8.2
Axillary LN
Inguinal LN
Lung
Liver
E
Tumor volume (mm3)
500
400
300
200
100
0
shControl
shSox4
*
**
Days
7
10
13
20
27
Tumor weight (mg)
400
300
200
100
0
shControl
shSox4


A
NMuMG
Empty vector
HA-Sox4
Fold enrichment
4.0
3.0
2.0
1.0
0.0
*
**
Interg
Ezh2
Interg
Ezh2
- TGFβ
+ TGFβ
B
Py2T
Empty vector
HA-Sox4
Fold enrichment
5.0
4.0
3.0
2.0
1.0
0.0
**
**
Interg
Ezh2
Interg
Ezh2
- TGFβ
+ TGFβ
C
NMuMG
Py2T
Ezh2 expression (fold change)
1.0
0.8
0.6
0.4
0.2
0.0
**
**
siControl
siSox4
siControl
siSox4
H3K27me3
H3
D
NMuMG
Relative luminescence
2.5
2.0
1.5
1.0
0.5
0.0
luc
Ezh2prom-luc
TGFβ
-
+
-
+
siRNA
siControl
siSox4
E
NMuMG
Sox4
Ezh2
mRNA expression (fold change)
3.0
2.0
1.0
0.0
***
***
**
**
***
*
siControl
+
-
+
-
+
-
+
-
siSox4
-
+
-
+
-
+
-
+
TGFβ (hrs)
0
0
2
2
4
4
8
8
F
Py2T
Sox4
Ezh2
mRNA expression (fold change)
3.0
2.0
1.0
0.0
**
***
**
*
**
*
*
siControl
+
-
+
-
+
-
+
-
siSox4
-
+
-
+
-
+
-
+
TGFβ (hrs)
0
0
2
2
4
4
8
8


A
NMuMG
TGFβ
0 day
1 day
4 days
7 days
siControl
siEzh2
siEzh1/2
B
NMuMG
siControl + - + - + - + -
siEzh2 - + - + - + - +
TGFβ 0 0 1 1 4 4 7 7
E-cadherin
N-cadherin
ZO-1
Ezh2
Actin
C
NMuMG
DAPI
Ecad
Ncad
Fn1
Merge
DAPI
ZO-1
Phalloidin
Paxillin
Merge
siControl
siEzh2
siEzh1/2
D
NMuMG
siControl
siEzh2
% Annexin V-positive cells
Days 0 1 4 7
TGFβ
E
NMuMG
G0/G1
S-G2/M
% of cells
siControl-d0
siEzh2-d0
siControl-d1
siEzh2-d1
siControl-d4
siEzh2-d4
F
NMuMG
siControl
siEzh2
# of cells (1x10³)
Days 2 4 7 10
TGFβ
G
Distant metastasis-free survival
Probability
low Sox4/Ezh2
highSox4/Ezh2
p = 0.000084
Time (days)


A
NMuMG
TGFβ
siControl
siControl + Ezh2
siSox4
siSox4 + Ezh2
0 day
1 day
4 days
B
Py2T
E-cadherin expression (fold change)
1.5
1.0
0.5
0.0
siControl
siSox4
Ezh2
TGFβ (days)
0
1
4
7
C
Py2T
Fibronectin expression (fold change)
20
15
10
5
0
siControl
siSox4
Ezh2
TGFβ (days)
0
1
4
7


A
301 genes
970 regions
Correlation coefficient between two replicates
Mean variance (in the two replicates)
B
NMuMG
Mcam
Pdgfrb
Sgsh
Fold enrichment above background
TGFβ
C
NMuMG
Fold enrichment above background
Itga5
Col1a1
St6galnac4
Gapdh
D
NMuMG
siControl
siEzh1/2
Fold enrichment above background
Mcam
Pdgfrb
Itga5
Gapdh
E
NMuMG
Mcam
siControl
siEzh1/2
Fold change
days
TGFβ
F
NMuMG
Pdgfrb
siControl
siEzh1/2
Fold change
days
TGFβ
G
NMuMG
Sgsh
siControl
siEzh1/2
Fold change
days
TGFβ
H
NMuMG
siControl
siEzh1/2
Fold change
Col4a3
Nfatc2
St6galnac6
Tns1